National Cancer Institute

Harnessing the Genomic Revolution: Investigation of the Effects of Radiation on Thyroid Cancer & Transgenerational Mutational Patterns

Stephen Chanock M.D.

Director, Division of Cancer Epidemiology & Genetics

National Cancer Institute

Beebe Symposium

November 2, 2016

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

NATIONAL CANCER INSTITUTE
Division of Cancer Epidemiology & Genetics

Acknowledgements

DCEG Leadership
Amy Berrington
Joseph Fraumeni
Robert Hoover
Peggy Tucker
Martha Linet

DCEG Chernobyl Unit
Kiyo Mabuchi
Alina Brenner
Maureen Hatch
Lisa Cahoon
Vladimir Drozdovitch
Mark Little
Andre Bouville

Endocrinology and Metabolism Institute, National Academy of Medical Sciences of Ukraine
Professor Mykola Tronko
Dr. Tetiana Bogdanova

National Research Centre for Radiation Medicine, Ukraine
Dr. Dimitry Bazyka
Vadim Chumak

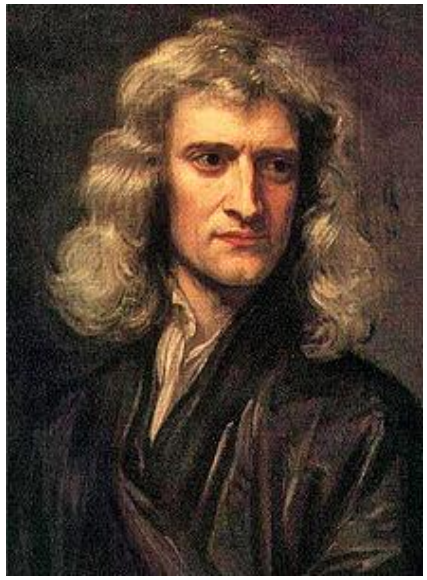
Chernobyl Tissue Bank
Professor Gerry Thomas
Dr. Angela Galpine

Cancer Genomics Research Laboratory
Meredith Yeager
Joe Boland
Amy Hutchinson
Belynda Hicks
Michael Dean
Michael Cullen
Sarah Wagner
Mia Steinberg

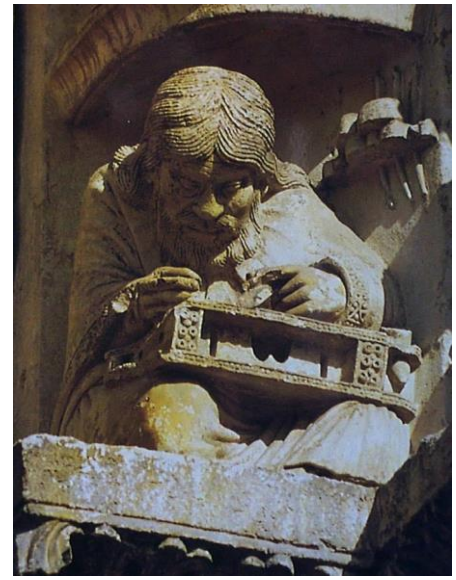
Broad Cancer Genome Analysis Group
Chip Stewart
Jaegil Kim
Paz Polack
Gad Getz

Beebe Symposium:

“Standing on the Shoulders of Giants”

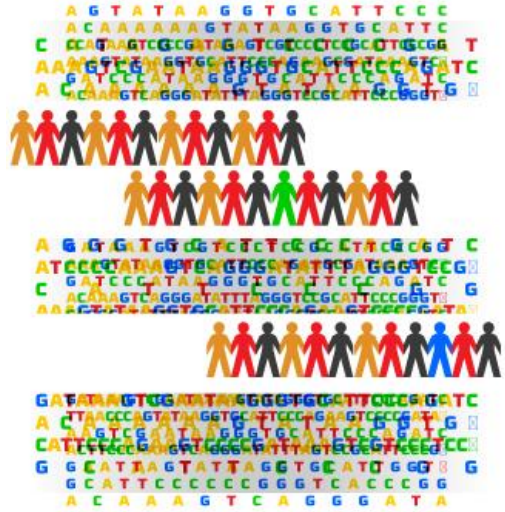


Sir Isaac Newton
1676

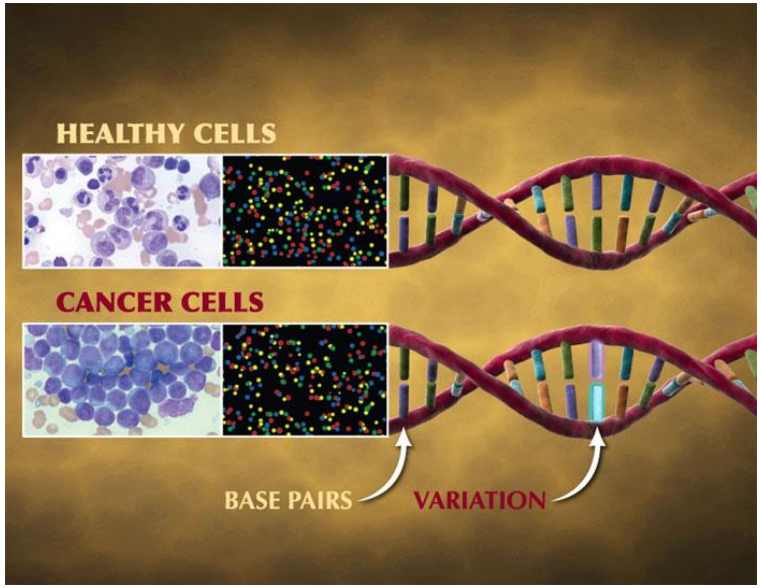


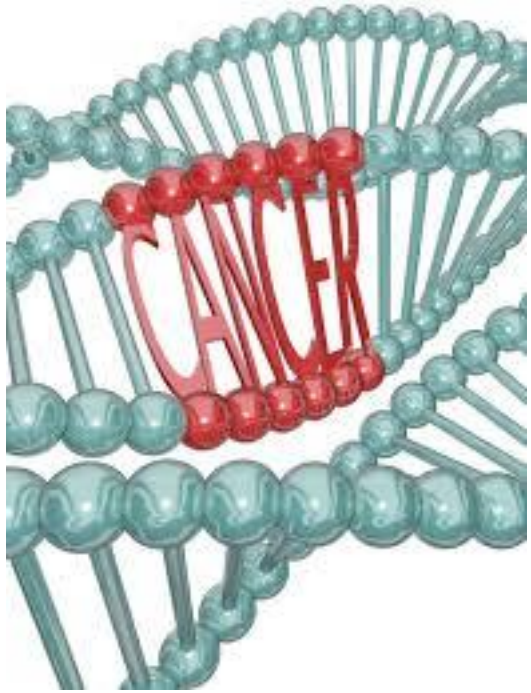
Bernard of Chartres
12th Century

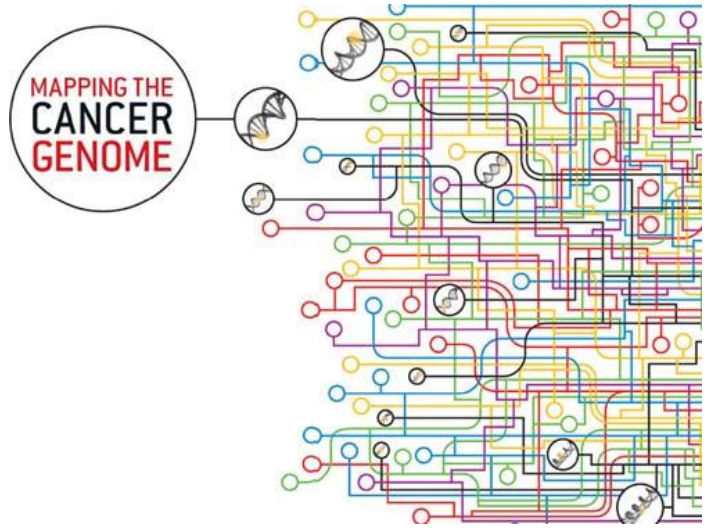
Harnessing the Genomic Revolution: 21st Century Investigation

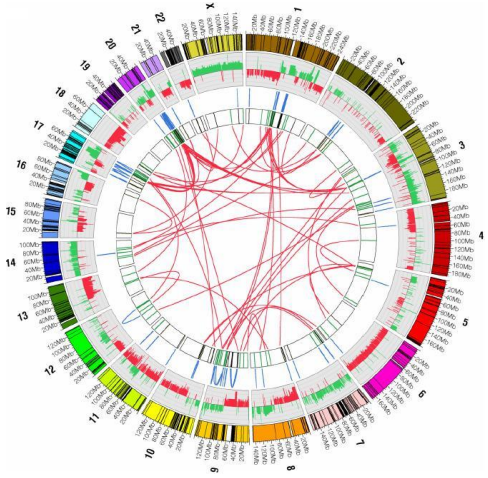


Understanding Cancer Genomes: Comprehensive Characterization of Errors









How and why do the errors occur?

Interaction with Environment- e.g., Radiation

DNA Replication- Background Error Rate

National Cancer Institute

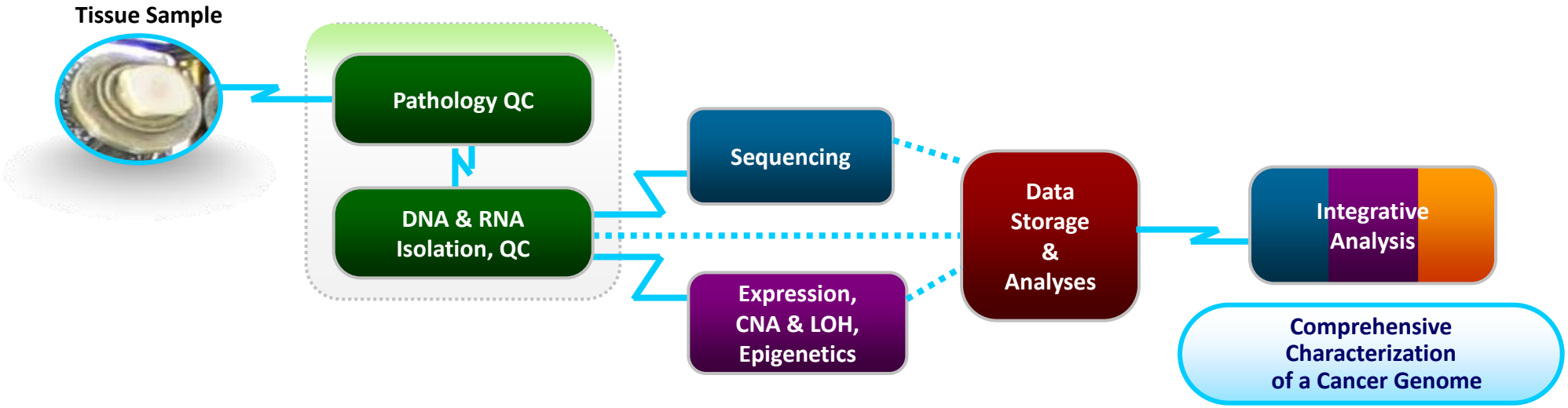
Full Genomic Characterization (TCGA-style) of Radiation-Related Thyroid Cancer in the Ukraine

Dr. Mykola Tronko
Director, Endocrinology and Metabolism Institute
National Academy of Medical Sciences of Ukraine

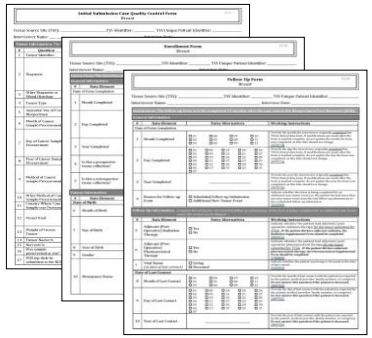
Head, Laboratory of Morphology of Endocrine System
VP Komissarenko Institute of Endocrinology & Metabolism

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

Model Pipeline for Comprehensive Characterization: The Cancer Genome Atlas (TCGA) for 20 Cancers



Chernobyl Thyroid Studies- CTB



SNP 6.0 ~45d

Methylation ~60d

miRNAseq ~105d

mRNAseq ~120d

DNAseq Exome/WGS ~180d

5 Major Analyses

1. Copy Number
2. Methylation
3. microRNAs
4. Large RNAs
5. Whole Genome/ Exome

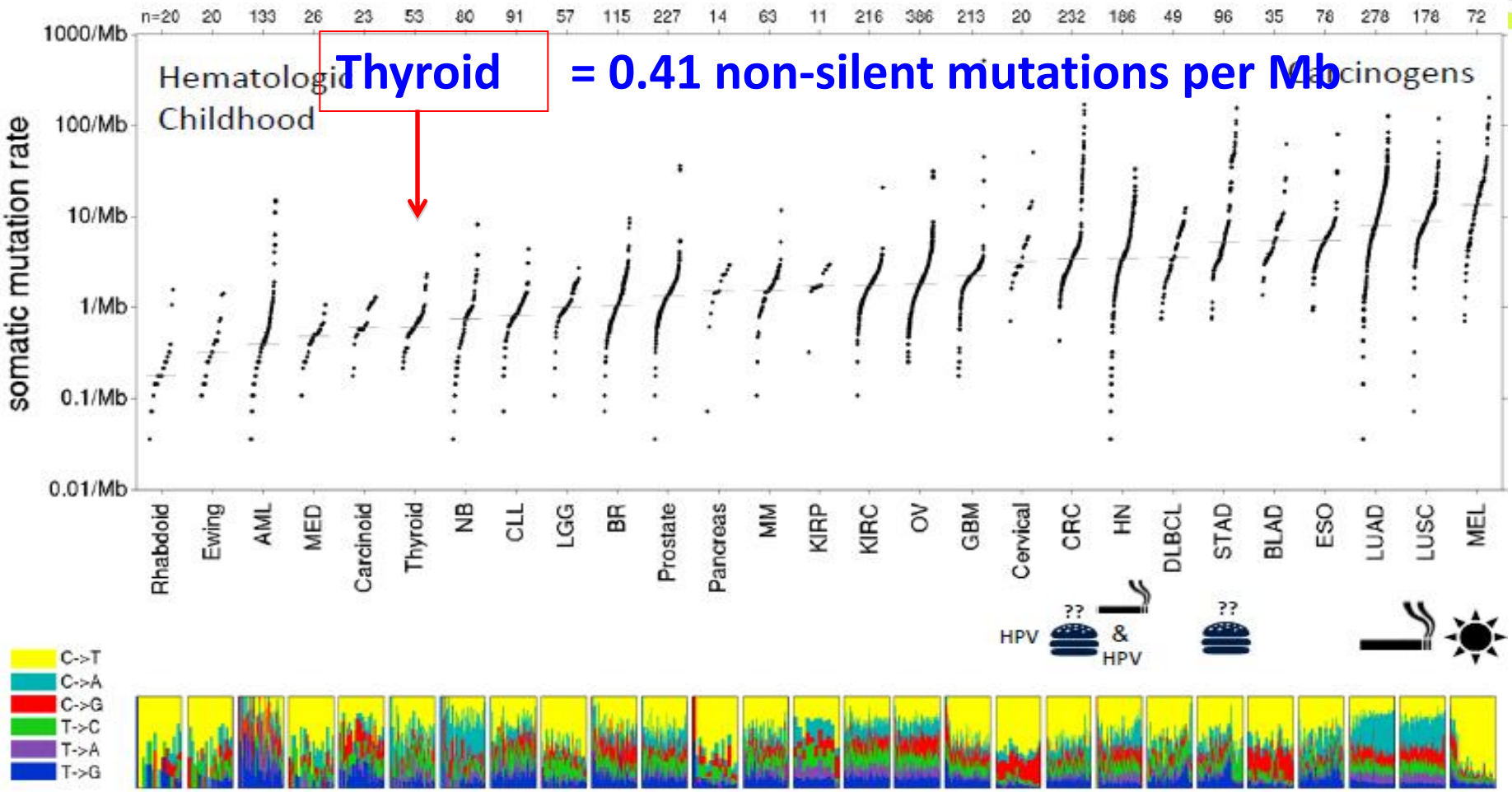
Goals of Genomic Characterization of Thyroid Cancers post Chernobyl

- Sample Size (fresh frozen tissue samples)
  - 450 radiation-related PTC cases from CTB
    - 100 UkrAm* & 350 non-UkrAm*
  - 50 sporadic (non-irradiated) PTC cases
    - CTB* cases born after Jan 1, 1987

- Analysis of "Paired" Samples
  - Dyads- Tumor plus normal- tissue or peripheral blood
  - Triads- Tumor plus non-cancer tissue plus peripheral blood
    - Opportunity to evaluate "field effect on non-cancerous tissue"
- *Currently- 334 cases are in production*

*From Zhytomyr, Chernihiv, Kyiv region, and Kyiv city

Lessons Learned from the Data
The Cancer Genome Atlas (TCGA)



M. Lawrence Nature 2013

The Cancer Genome Atlas

Resource

Cell

Integrated Genomic Characterization of Papillary Thyroid Carcinoma

The Cancer Genome Atlas Research Network[1,*]

[1]The Cancer Genome Atlas Program Office, National Cancer Institute at NIH, 31 Center Drive, Bldg. 31, Suite 3A20, Bethesda, MD 20892, USA

*Correspondence: giordano@umich.edu, gadgetz@broadinstitute.org

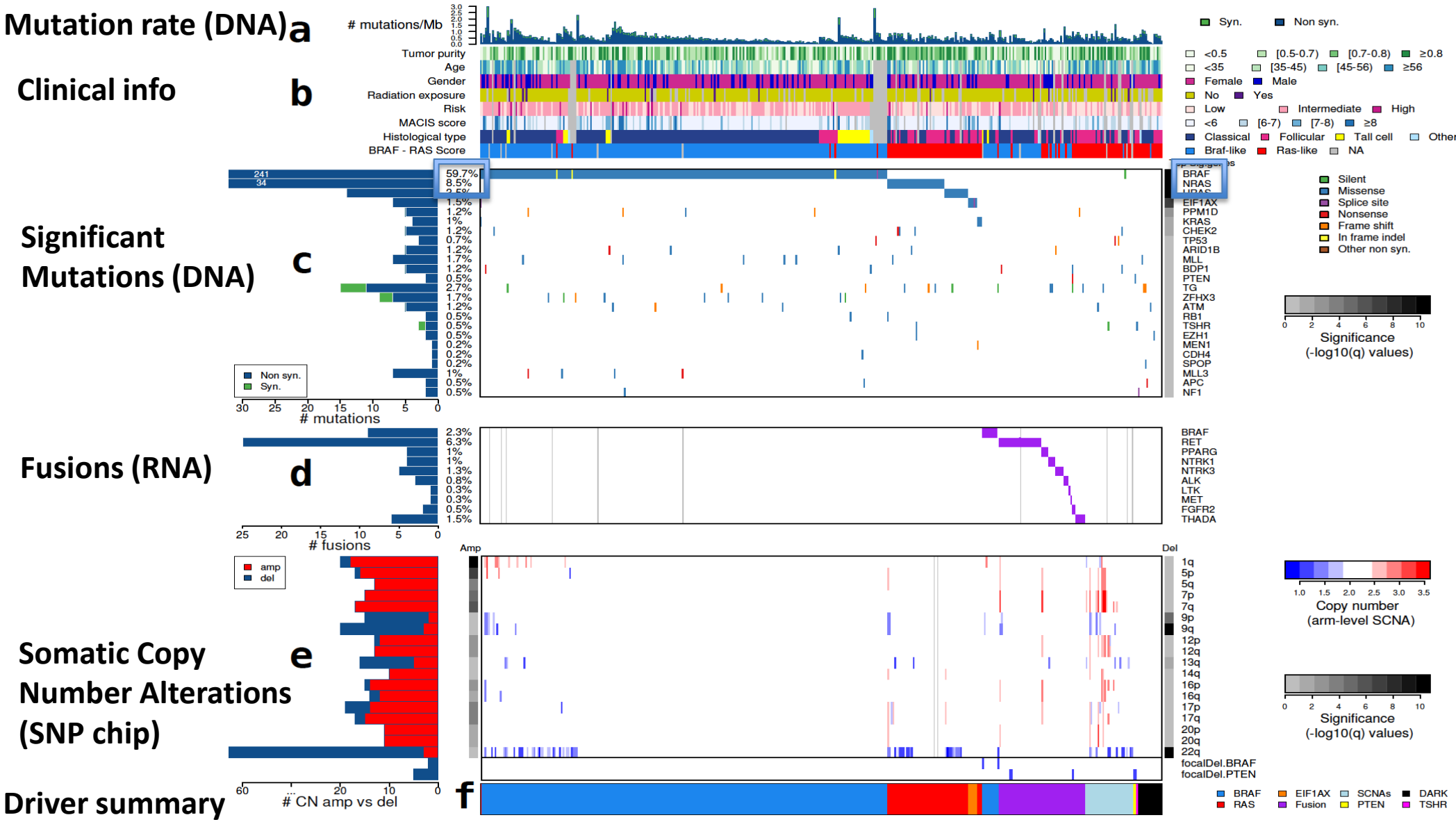
http://dx.doi.org/10.1016/j.cell.2014.09.050

SUMMARY

Papillary thyroid carcinoma (PTC) is the most common type of thyroid cancer. Here, we describe the genomic landscape of 496 PTCs. We observed a low frequency of somatic alterations (relative to other carcinomas) and extended the set of known PTC driver alterations to include *EIF1AX*, *PPM1D*, and *CHEK2* and diverse gene fusions. These discoveries reduced the fraction of PTC cases with unknown oncogenic driver from 25% to 3.5%. Combined

Previous genetic studies report a high frequency (70%) of activating somatic alterations of genes encoding effectors in the mitogen-activated protein kinase (MAPK) signaling pathway, including point mutations of *BRAF* and the *RAS* genes (Cohen et al., 2003; Kimura et al., 2003; Lemoine et al., 1988; Suárez et al., 1988), as well as fusions involving the *RET* (Grieco et al., 1990) and *NTRK1* tyrosine kinases (Pierotti et al., 1995). These mutations are almost always mutually exclusive (Soares et al., 2003), suggesting similar or redundant downstream effects. The various MAPK pathway alterations are strongly associated with distinct clinicopathological characteristics (Adeniran et al., 2006), and gene expression (Giordano et al., 2005) and DNA

Overview of Somatic Alterations in TCGA Thyroid Cancer



Preliminary Assessments of Somatic Mutations in Radiation-induced Thyroid Cancer: *Are There Signatures or Patterns?*

Opportunity for a Large-Scale 'Agnostic' Examination

Journal List › J Clin Endocrinol Metab › PMC3913801

THE ENDOCRINE SOCIETY JCEM About JCEM Submission Subscription

J Clin Endocrinol Metab. 2014 Feb; 99(2): E276–E285. PMCID: PMC3913801
Published online 2013 Nov 18. doi: 10.1210/jc.2013-2503

The Increase in Thyroid Cancer Incidence During the Last Four Decades Is Accompanied by a High Frequency of *BRAF* Mutations and a Sharp Increase in *RAS* Mutations

Chan Kwon Jung, Mark P. Little, Jay H. Lubin, Alina V. Brenner, Samuel A. Wells, Jr., Alice J. Sigurdson, and Yuri E. Nikiforov

Author information ► Article notes ► Copyright and License information ►

Abstract Go to:

Context:

Thyroid cancer incidence rates in the United States and globally have increased steadily over the last 40 years, primarily due to a tripling of the incidence of papillary thyroid carcinoma (PTC).

Objective:

The purpose of this study was to analyze trends in demographic, clinical, pathologic, and molecular characteristics of PTC from 1974 to 2009.

Design and Setting:

We identified and histologically reviewed 469 consecutive cases of PTC from one US institution from 4 preselected periods (1974 to 1985, 1990 to 1992, 2000, and 2009) and assessed *BRAF* and *RAS* point mutations and *RET/PTC* rearrangements among 341 tumors ≥0.3 cm in size. Changes over time were analyzed using polytomous and binary logistic regression; all analyses were adjusted for age and sex.

Results:

During this period, the median age of patients at diagnosis increased from 37 to 53 years ($P < .001$) and the

Carcinogenesis Advance Access published September 10, 2015

OXFORD

Carcinogenesis, 2015, 1–7

doi:10.1093/carcin/bgv119
Original Manuscript

ORIGINAL MANUSCRIPT

Genomic copy number analysis of Chernobyl papillary thyroid carcinoma in the Ukrainian–American Cohort

Martin Selmansberger, Herbert Braselmann, Julia Hess, Tetiana Bogdanova[1], Michael Abend[2], Mykola Tronko[1], Alina Brenner[3], Horst Zitzelsberger, and Kristian Unger*

Research Unit Radiation Cytogenetics, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, 85764 Neuherberg, Germany, [1]Institute of Endocrinology and Metabolism, National Academy of Medical Sciences of the Ukraine, 254114 Kiev, Ukraine, [2]Bundeswehr Institute of Radiobiology, 80937 Munich, Germany and [3]Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, NCI, Bethesda, MD 20892, USA

*To whom correspondence should be addressed. Tel: +49 89 3187 3515; Fax: +49 09 3187 2873;
Email: unger@helmholtz-muenchen.de

Abstract

One of the major consequences of the 1986 Chernobyl reactor accident was a dramatic increase in papillary thyroid carcinoma (PTC) incidence, predominantly in patients exposed to the radioiodine fallout at young age. The present study is the first on genomic copy number alterations (CNAs) of PTCs of the Ukrainian–American cohort (UkrAm) generated by array comparative genomic hybridization (aCGH). Unsupervised hierarchical clustering of CNA profiles revealed a significant enrichment of a subgroup of patients with female gender, long latency (>17 years) and negative lymph node status. Further, we identified single CNAs that were significantly associated with latency, gender, radiation dose and BRAF V600E mutation status. Multivariate analysis revealed no interactions but additive effects of parameters gender, latency and dose on CNAs. The previously identified radiation-associated gain of the chromosomal bands 7q11.22–11.23 was present in 29% of cases. Moreover, comparison of our radiation-associated PTC data set with the TCGA data set on sporadic PTCs revealed altered copy numbers of the tumor driver genes NF2 and CHEK2. Further, we integrated the CNA data with transcriptomic data that were available on a subset of the herein analyzed cohort and did not find statistically significant associations between the two molecular layers. However, applying hierarchical clustering on a 'BRAF-like/RAS-like' transcriptome signature split the cases into four groups, one of which containing all BRAF-positive cases validating the signature in an independent data set.

Introduction

One of the major consequences of the 1986 Chernobyl nuclear accident to human health was a dramatic increase in thyroid cancer incidence among those who were children or adolescents at the time of exposure (1,2). Several studies clearly related this increase to radioiodine exposure (mainly ^{131}I) from the reactor accident (3–7). Numerous genomic studies on post-Chernobyl papillary thyroid carcinomas (PTCs) have been published so far. A study by Unger *et al.* (8) has shown intratumoral heterogeneity of RET/PTC rearrangements in post-Chernobyl PTCs suggesting either a multiclonal origin of the tumors or RET/PTC being a late subclonal event. A comparative study on global genomic copy number alterations (CNAs) in two age-matched cohorts of post-Chernobyl PTCs and sporadic PTCs revealed a radiation-specific genomic copy number gain of the chromosomal bands 7q11.22–11.23, which was exclusively detected in exposed cases by array comparative genomic hybridization (aCGH) (9). Further, a significant mRNA overexpression of the CLIP2 gene, located in the gained band, was shown (9). Among others, the largest cohort of subjects who were exposed to the radioiodine-contaminated fallout from the Chernobyl accident under the age of 18 is the so called Ukrainian–American (UkrAm) cohort (10), which was established by the US National Cancer Institute and the

Received: March 19, 2015; Revised: July 6, 2015; Accepted: August 13, 2015

Original Article

RET/PTC and *PAX8/PPARγ* Chromosomal Rearrangements in Post-Chernobyl Thyroid Cancer and Their Association With Iodine-131 Radiation Dose and Other Characteristics

Rebecca J. Leeman-Neill, MD, PhD[1]; Alina V. Brenner, MD, PhD, MPH[2]; Mark P. Little, MA, DPhil[2]; Tetiana I. Bogdanova, PhD[3]; Maureen Hatch, PhD[2]; Liudmyla Y. Zurnadzy, MD, PhD[3]; Kiyohiko Mabuchi, MD, DrPH[2]; Mykola D. Tronko, MD, PhD[3]; and Yuri E. Nikiforov, MD, PhD[1]

BACKGROUND: Childhood exposure to iodine-131 from the 1986 nuclear accident in Chernobyl, Ukraine, led to a sharp increase in papillary thyroid carcinoma (PTC) incidence in regions surrounding the reactor. Data concerning the association between genetic mutations in PTCs and individual radiation doses are limited. **METHODS:** Mutational analysis was performed on 62 PTCs diagnosed in a Ukrainian cohort of patients who were < 18 years old in 1986 and received 0.008 to 8.6 Gy of ^{131}I to the thyroid. Associations between mutation types and ^{131}I dose and other characteristics were explored. **RESULTS:** *RET/PTC* (ret proto-oncogene/papillary thyroid carcinoma) rearrangements were most common (35%), followed by *BRAF* (15%) and *RAS* (8%) point mutations. Two tumors carrying *PAX8/PPARγ* (paired box 8/peroxisome proliferator-activated receptor gamma) rearrangement were identified. A significant negative association with ^{131}I dose for *BRAF* and *RAS* point mutations and a significant concave association with ^{131}I dose, with an inflection point at 1.6 Gy and odds ratio of 2.1, based on a linear–quadratic model for *RET/PTC* and *PAX8/PPARγ* rearrangements were found. The trends with dose were significantly different between tumors with point mutations and rearrangements. Compared with point mutations, rearrangements were associated with residence in the relatively iodine-deficient Zhytomyr region, younger age at exposure or surgery, and male sex. **CONCLUSIONS:** These results provide the first demonstration of *PAX8/PPARγ* rearrangements in post-Chernobyl tumors and show different associations for point mutations and chromosomal rearrangements with ^{131}I dose and other factors. These data support the relationship between chromosomal rearrangements, but not point mutations, and ^{131}I exposure and point to a possible role of iodine deficiency in generation of *RET/PTC* rearrangements in these patients. *Cancer* 2013;119:1792–9. *© 2013 American Cancer Society.*

KEYWORDS: Chernobyl, papillary thyroid carcinoma, iodine-131, *RET/PTC*, *PAX8/PPARγ*.

INTRODUCTION

Exposure to ionizing radiation during childhood is known to cause thyroid cancer, with a significantly dose-dependent increased incidence, particularly in children and young adults.[1] After the nuclear accident in April 1986 in Chernobyl, Ukraine, residents of regions surrounding the Chernobyl nuclear power plant, including Ukraine, Belarus, and the Russian Federation, received variable doses of radioiodines through inhalation and ingestion of contaminated dairy products or vegetables. These regions experienced a dramatic rise in incidence of thyroid cancer,[2] with at least 5000 new cases observed in individuals exposed during childhood or adolescence.[3] Papillary thyroid carcinoma (PTC) is known to be the principal type of thyroid carcinoma associated with radiation exposure and comprised the majority of pediatric thyroid tumors in residents of the regions surrounding Chernobyl.[2,4,5]

Case-control and cohort studies of post-Chernobyl thyroid cancers have demonstrated that the risk of thyroid carcinoma is strongly related to ^{131}I dose absorbed by the thyroid.[6–9] The reported excess relative risk per unit of dose (Gy) is between 2 and 5. In addition, age at exposure and iodine deficiency have been found to modify the ^{131}I-related risk of thyroid cancer, with higher risk per unit of dose observed in persons exposed as younger children, particularly infants,[10] and in individuals living in areas with low soil iodine content.[7]

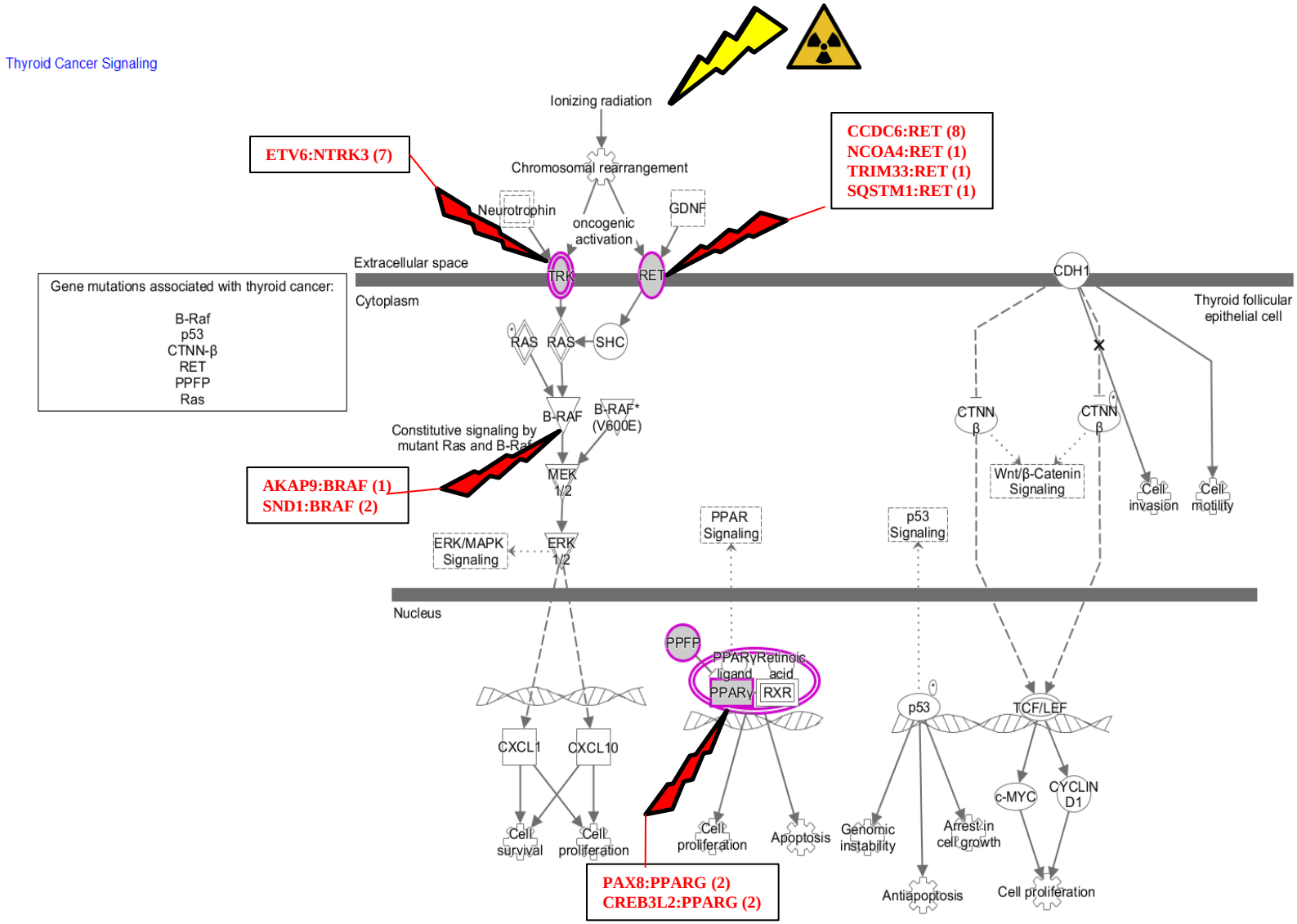
Corresponding author: Yuri Nikiforov, MD, PhD, Department of Pathology, University of Pittsburgh, 200 Lothrop Street, Presbyterian University Hospital, Room C-606, Pittsburgh, PA 15213; Fax: (412) 802-6799; nikiforovye@upmc.edu

[1]Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania; [2]Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland; [3]State Institution V.P. Komisarenko Institute of Endocrinology and Metabolism of the National Academy of Medical Sciences of Ukraine, Kyiv, Ukraine

It is with great sadness that we report the death of our colleague Dr. Elaine Ron, who was one of the original investigators involved in design of the study. We greatly appreciate her contribution and support. We thank Dr. Gerry Thomas and staff of the Chernobyl Tissue Bank for providing samples. We are grateful to the dosimetry team, including Drs. Ilya Likhtarev, Lina Kovgan (Radiation Protection Institute, Ukrainian Academy of Technological Sciences), Andre Bouville (formerly with the National Cancer Institute [NCI], National Institutes of Health [NIH]), and Vladimir Drozdovich (NCI, NIH) for dose reconstruction efforts. We thank Dr. Alice Sigurdson (NCI, NIH) for critical review of the manuscript.

DOI: 10.1002/cncr.27893, **Received:** August 21, 2012; **Revised:** September 25, 2012; **Accepted:** October 4, 2012, **Published online** February 21, 2013 in Wiley Online Library (wileyonlinelibrary.com)

Central Thyroid Cancer Signaling Pathway:



What is a Mutation Signature?

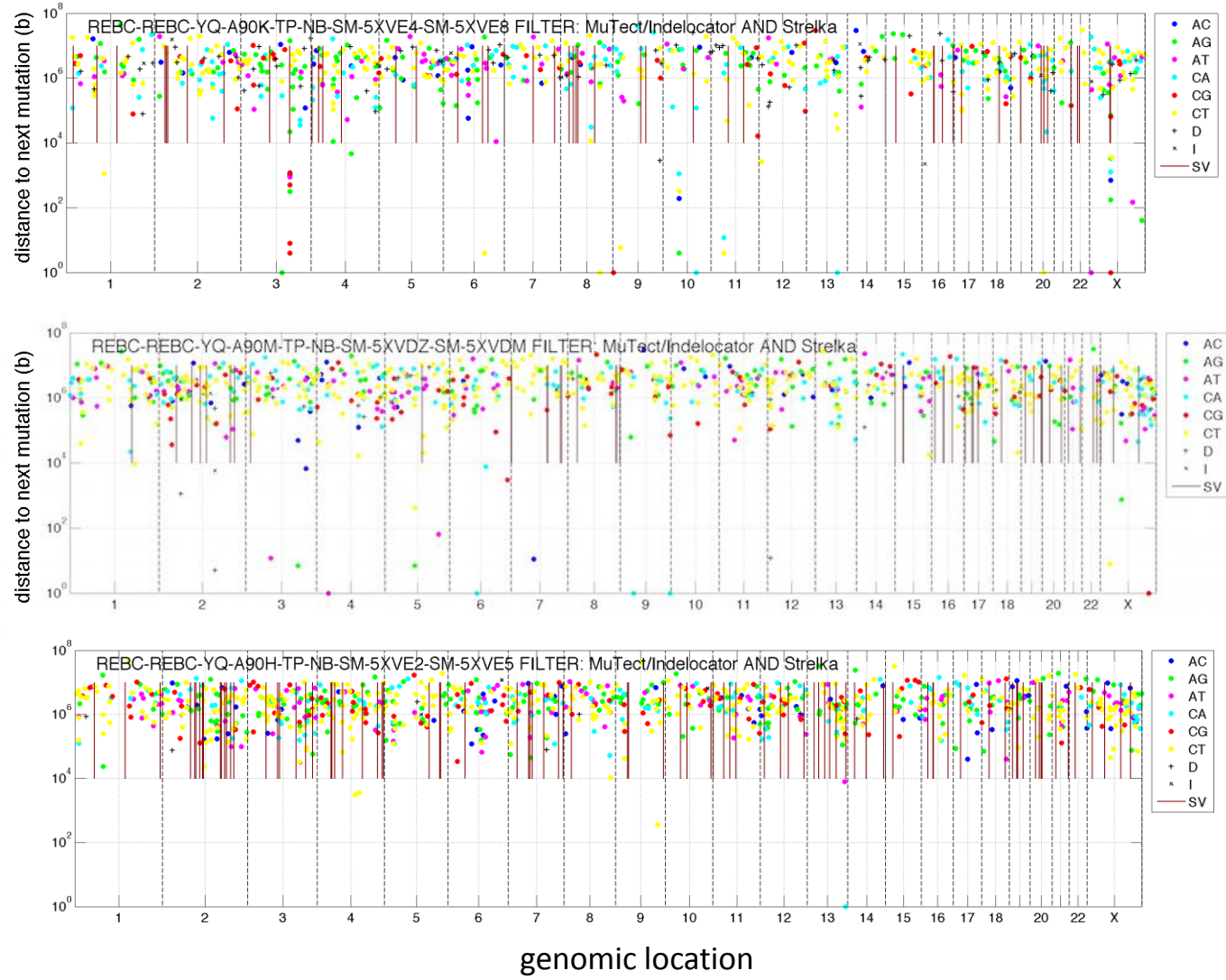
Somatic mutations are present in all cells of the human body and occur throughout life. They are the consequence of multiple mutational processes, including the intrinsic slight infidelity of the DNA replication machinery, exogenous or endogenous mutagen exposures, enzymatic modification of DNA and defective DNA repair. Different mutational processes generate unique combinations of mutation types, termed "Mutational Signatures".

COSMIC DEFINITION

Alexandrov Nature 2013

Mutation Genomic Profiles:
Little evidence for Kategis- ‘concetrated mutations’

Due to the exceptionally low mutation density, clusters of mutations (kategis) are less likely.



Somatic Copy Number Alterations
35 of 197 tumors (SNP Chip)

Chr	STATE	Counts	Subject Counts	Freq	genes in the region	reported at TCGA Cell paper
1	GAIN	1	1	0.51%		Yes
2	GAIN	1	1	0.51%		
3	GAIN	1	1	0.51%		
4	GAIN	3	3	1.52%		
5	GAIN	2	2	1.02%		yes
6	GAIN	3	3	1.52%	Oncogene: DEK, E2F3?	
7	GAIN	2	2	1.02%		yes
8	GAIN	2	2	1.02%		
9	GAIN	1	1	0.51%		yes
10	GAIN	1	1	0.51%		
11	GAIN	1	1	0.51%		
12	GAIN	1	1	0.51%		yes
13q	GAIN	1	1	0.51%		yes
14q	GAIN	3	3	1.52%		yes
15q	GAIN	1	1	0.51%		
16	GAIN	2	2	1.02%		yes
17	GAIN	4	4	2.03%		yes
20	GAIN	1	1	0.51%		yes
22q	GAIN	4	4	2.03%		
2	**LOSS**	**2**	**2**	**1.02%**	**ALK-STRN fusion (0.2% in TCGA)**	yes
5	LOSS	1	1	0.51%	**Tumor suppressor: APC, MCC**	
6	LOSS	1	1	0.51%		
9	LOSS	1	1	0.51%		yes
10	LOSS	1	1	0.51%	not contain PTEN	
13q	LOSS	3	3	1.52%	**Tumor suppressor: RB1**	yes
15q	LOSS	2	2	1.02%	no	
16	LOSS	3	3	1.52%	**Tumor suppressor: CDH1, CDH11, CBFA2T3**	yes
22q	**LOSS**	**19**	**18**	**9.14%**	**Tumor suppressor: CHEK2, NF2 (14.1% in TCGA); enrich for follicular variant (FV) subtype**	yes
21q	NEUTRAL	1	1	0.51%		
22q	NEUTRAL	1	1	0.51%		
Total		**70**				

22q deletion

Lower than TCGA

Interim Conclusions

- Very Quiet Somatic Profile of Mutations
- Likely Signature of Ageing
- BRAF mutations in older individuals with lower Radiation Doses
- An Increase in Fusion Drivers against a minimal background mutational spectra

National Cancer Institute

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

Parental Irradiation of Ukrainian Clean-up Workers and Evacuees and Germline Mutations in their Offspring (TRIO Study)

Dr. Dimitry Bazyka
Director-General, National Research Centre for Radiation Medicine
Academy of Medical Science of Ukraine

Background

- Little evidence of untoward pregnancy outcomes, childhood mortality, or sex chromosome aneuploidy associated with parental radiation exposure in Japanese A-bomb F_1 or other studies
- Based on 7-locus mouse data (Russell et al) and non-significant indications from the A-bomb F_1 study, ICRP assumes parental radiation exposure induces a large spectrum of genetic effects in offspring
  - Doubling dose (DD) of approximately 1 Gy

DD = Radiation dose expected to double the spontaneous mutations rate in a generation

Trio Analyses of Mutational Patterns

Parent-of-origin-specific signatures of *de novo* mutations

Jakob M Goldmann[1,10], Wendy S W Wong[2,10], Michele Pinelli[3], Terry Farrah[4], Dale Bodian[2], Anna B Stittrich[4], Gustavo Glusman[4], Lisenka E L M Vissers[5], Alexander Hoischen[5], Jared C Roach[4], Joseph G Vockley[2,6], Joris A Veltman[5,7], Benjamin D Solomon[2,8,9], Christian Gilissen[5,11] & John E Niederhuber[2,9,11]

***De novo* mutations (DNMs) originating in gametogenesis are an important source of genetic variation. We use a data set of 7,216 autosomal DNMs with resolved parent of origin from whole-genome sequencing of 816 parent–offspring trios to investigate differences between maternally and paternally derived DNMs and study the underlying mutational mechanisms. Our results show that the number of DNMs in offspring increases not only with paternal age, but also with maternal age, and that some genome regions show enrichment for maternally derived DNMs. We identify parent-of-origin-specific mutation signatures that become more pronounced with increased parental age, pointing to different mutational mechanisms in spermatogenesis and oogenesis. Moreover, we find DNMs that are spatially clustered to have a unique mutational signature with no significant differences between parental alleles, suggesting a different mutational mechanism. Our findings provide insights into the molecular mechanisms that underlie mutagenesis and are relevant to disease and evolution in humans[1].**

Studies of *de novo* mutations (DNMs) in humans have estimated the mutation rate of single-nucleotide variants to be approximately 1×10^{-8} mutations per generation, giving rise to 45–60 DNMs per genome[2–5]. The susceptibility to DNMs varies by several orders of magnitude along the genome and may be influenced by factors such as nucleotide content, replication timing, distance to recombination hotspots, nucleosome occupancy, transcription, and chromatin 'openness'[4,6]. Several mechanisms of DNA mutation are known, most predominantly involving DNA replication[7]. The latter mechanism also explains the 3.9:1 ratio of DNMs on the paternal allele to the maternal allele, as there are many more germline cell divisions in spermatogenesis than in oogenesis[2]. We hypothesize that the different underlying biology of male and female gametogenesis results in differences in mutational signatures between paternally and maternally transmitted DNMs. These signatures will provide insight into the mechanisms underlying *de novo* mutations in human germline cells.

Studies to date have lacked sufficient sample size to determine the parental allele for large numbers of DNMs so as to compare DNMs of paternal and maternal origin. In this study, whole-genome sequencing (WGS) was performed on 832 offspring–parent trios, with an average of 60× coverage, by Complete Genomics Inc. (**Table 1**, **Supplementary Tables 1–4**; see Online Methods for a description of the cohort)[8]. After removing an outlier and one twin from each of the monozygotic twin pairs, *de novo* mutations were identified for the autosomes of 816 trios. A random forest classifier was used to remove potential false positives from the initial set of putative DNMs, resulting in 36,441 DNMs, or an average of 45 DNMs per individual (Online Methods and **Supplementary Tables 5–8**). Quality assessment of these results based on monozygotic twin concordance and Sanger validations of a subset of DNMs indicated high specificity (**Supplementary Tables 9–11**, Online Methods). Overall, the nucleotide substitution frequencies for DNMs were dominated by C–T and T–C changes, giving rise to a transition/transversion ratio (Ts/Tv) of 2.23. Haplotype assembly of all mutations successfully phased 19.8% of all DNMs, resulting in a set of 7,216 phased DNMs (Online Methods, **Supplementary Tables 12–15**). Assessing the parental origin of DNMs, we found that 5,640 DNMs were on the paternal allele and 1,576 on the maternal allele, giving rise to the expected median paternal/maternal ratio of 3.6:1 (**Supplementary Fig. 1**)[2,3].

Multiple studies have shown that the numbers of DNMs in offspring are positively correlated with increasing paternal age at the time of conception[2,3]. Using our phased DNMs, we were able to confirm

Table 1 Cohort description

Birth constellation	No. births	No. children	No. sequenced samples
Singletons	731	731	2,193
Dizygotic twins	35	70	140
Monozygotic twins	14	28	56
Triplet	1	3	5
Total	781	832	2,394

The cohort consists of 731 trios, 49 quartets, and one quintet, resulting in a total of 832 children.

[1]Department of Human Genetics, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, Nijmegen, the Netherlands. [2]Inova Translational Medicine Institute (ITMI), Inova Health Systems, Falls Church, Virginia, USA. [3]Telethon Institute of Genetics and Medicine (TIGEM), Naples, Italy. [4]Institute for Systems Biology, Seattle, Washington, USA. [5]Department of Human Genetics, Donders Centre for Neuroscience, Radboud University Medical Center, Nijmegen, the Netherlands. [6]Department of Pediatrics, Virginia Commonwealth University School of Medicine, Richmond, Virginia, USA. [7]Department of Clinical Genetics, GROW School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, the Netherlands. [8]Department of Pediatrics, Inova Children's Hospital, Inova Health System, Falls Church, Virginia, USA. [9]Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [10]These authors contributed equally to this work. [11]These authors jointly supervised this work. Correspondence should be addressed to C.G. (Christian.gilissen@radboudumc.nl) or J.E.N. (John.Niederhuber@inova.org).

Received 1 December 2015; accepted 26 May 2016; published online 20 June 2016; doi:10.1038/ng.3597

Large-scale whole-genome sequencing of the Icelandic population

Daniel F Gudbjartsson[1,2,21], Hannes Helgason[1,2,21], Sigurjon A Gudjonsson[1], Florian Zink[1], Asmundur Oddson[1], Arnaldur Gylfason[1], Soren Besenbacher[3], Gisli Magnusson[1], Bjarni V Halldorsson[1,4], Eirikur Hjartarson[1], Gunnar Th Sigurdsson[1], Simon N Stacey[1], Michael L Frigge[1], Hilma Holm[1,5], Jona Saemundsdottir[1], Hafdis Th Helgadottir[1], Hrefna Johannsdottir[1], Gunnlaugur Sigfusson[6], Gudmundur Thorgeirsson[7,8], Jon Th Sverrisson[9], Solveig Gretarsdottir[1], G Bragi Walters[1], Thorunn Rafnar[1], Bjarni Thjodleifsson[7], Einar S Bjornsson[8,10], Sigurdur Olafsson[8,10], Hildur Thorarinsdottir[10], Thora Steingrimsdottir[8,11], Thora S Gudmundsdottir[11], Asgeir Theodors[10], Jon G Jonasson[8,12,13], Asgeir Sigurdsson[1], Gyda Bjornsdottir[1], Jon J Jonsson[14,15], Olafur Thorarensen[16], Petur Ludvigsson[16], Hakon Gudbjartsson[1,2], Gudmundur I Eyjolfsson[17], Olof Sigurdardottir[18], Isleifur Olafsson[19], David O Arnar[7,8], Olafur Th Magnusson[1], Augustine Kong[1,2], Gisli Masson[1], Unnur Thorsteinsdottir[1,8], Agnar Helgason[1,20], Patrick Sulem[1] & Kari Stefansson[1,8]

Here we describe the insights gained from sequencing the whole genomes of 2,636 Icelanders to a median depth of 20×. We found 20 million SNPs and 1.5 million insertions-deletions (indels). We describe the density and frequency spectra of sequence variants in relation to their functional annotation, gene position, pathway and conservation score. We demonstrate an excess of homozygosity and rare protein-coding variants in Iceland. We imputed these variants into 104,220 individuals down to a minor allele frequency of 0.1% and found a recessive frameshift mutation in *MYL4* that causes early-onset atrial fibrillation, several mutations in *ABCB4* that increase risk of liver diseases and an intronic variant in *GNAS* associating with increased thyroid-stimulating hormone levels when maternally inherited. These data provide a study design that can be used to determine how variation in the sequence of the human genome gives rise to human diversity.

The advent of high-throughput genotyping and sequencing has revolutionized the ability to investigate how diversity in the sequence of the human genome affects human diversity[1]. Large-scale genotyping of common variants led to an avalanche of discoveries of variants associating with common and complex diseases[2]. Now studies based on whole-genome and exome sequencing are beginning to yield rare variants associating with common diseases[3–12]. They also provide unprecedented information about human sequence diversity and insights into the structure and history of human populations[13–16]. Several large-scale sequencing projects are ongoing or in the planning stages, foremost among them the 1000 Genomes Project[16] and the Exome Sequencing Project (ESP)[13,14], which have already provided valuable information about human genome diversity and tools to use in genetic discovery.

Our efforts at studying the human genome and its impact on diseases and other traits have focused on the Icelandic population. Genetic studies of the Icelandic population benefit from a genealogy of the nation reaching centuries back in time, a founder effect and broad access to nationwide healthcare information. The transition from genome-wide association studies (GWAS) based on common SNPs on microarrays to those based on a vast number of rare variants identified by whole-genome and exome sequencing presents new opportunities and challenges.

Here we describe the insights gained from sequencing the whole genomes of 2,636 Icelanders. First, we describe the density and

[1]deCODE Genetics/Amgen, Inc., Reykjavik, Iceland. [2]School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland. [3]Bioinformatics Research Centre, Aarhus University, C.F. Møllers Alle, Aarhus, Denmark. [4]Institute of Biomedical and Neural Engineering, Reykjavik University, Reykjavik, Iceland. [5]Division of Cardiovascular Diseases, Mayo Clinic, Rochester, Minnesota, USA. [6]Children's Hospital, Landspitali University Hospital, Reykjavik, Iceland. [7]Department of Medicine, Landspitali University Hospital, Reykjavik, Iceland. [8]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [9]Department of Internal Medicine, Akureyri Hospital, Akureyri, Iceland. [10]Department of Internal Medicine, Division of Gastroenterology and Hepatology, Landspitali University Hospital, Reykjavik, Iceland. [11]Department of Obstetrics and Gynecology, Landspitali University Hospital, Reykjavik, Iceland. [12]Department of Pathology, Landspitali University Hospital, Reykjavik, Iceland. [13]Icelandic Cancer Registry, Reykjavik, Iceland. [14]Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [15]Department of Genetics and Molecular Medicine, Landspitali University Hospital, Reykjavik, Iceland. [16]Department of Pediatrics, Section of Child Neurology, The Children's Hospital of Reykjavik, Landspitali University Hospital, Reykjavik, Iceland. [17]Icelandic Medical Center (Laeknasetrid), Laboratory in Mjodd (RAM), Reykjavik, Iceland. [18]Department of Clinical Biochemistry, Akureyri Hospital, Akureyri, Iceland. [19]Department of Clinical Biochemistry, Landspitali University Hospital, Reykjavik, Iceland. [20]Department of Anthropology, University of Iceland, Reykjavik, Iceland. [21]These authors contributed equally to this work. Correspondence should be addressed to D.F.G. (daniel.gudbjartsson@decode.is) or K.S. (kari.stefansson@decode.is).

Received 17 February 2014; accepted 13 February 2015; published online 25 March 2015; doi:10.1038/ng.3247

Paternal 80% & Maternal 20%

Recombination Events
? Increase with Radiation

Comprehensive characterization of germline genomes of trios (parents and children) with pre-conception exposure to radiation from the Chernobyl accident

- **Phase 1**
  - **Pilot: 48 Trios (mother/father and child)**
  - **All offspring born > 1year after accident**
  - **Collection of Whole Blood (informed consent)**
  - ***In progress***
    - ***SNP Chip completed (156 samples)***
    - ***WGS received***

Target Trio Numbers

- Initial study: Recruit 50 trios, selected from risk categories (10 trios for each of 5 groups):
  - Exposed father, exposed mother
  - Exposed father, unexposed mother
  - Unexposed father, exposed mother
  - Unexposed father, unexposed mother
  - High dose emergency worker (fathers only, with acute radiation syndrome)
- Full study aims to recruit up to 450 trios from exposed and/or unexposed parents

Analysis Plan

- **Germline Characterization**
  - **Whole Genome Sequencing 60X**
  - **SNP- microarray**
  - **Methyl-Microarray**
  - **Whole blood for RNA Sequencing (if indicated)**

- **Assess Rates of:**
  - **Minisatellite Mutations (early studies)**
  - ***de novo* Mutations (*not seen in parents*)- including rates**
  - **Copy Number/Structural Variations**
  - **Recombination Rates (across generations)**
  - **Somatic Mutations & Detectable Mosaicism**
  - **Variation in Telomere Length (RT-PCR + WGS)**
  - **Methylation Patterns**